آمار ریاضی ۱
فصل دوم: قضیه باسو

استاد: دکتر اکبر اصغرزاده

قضیه باسو $(Basu\ Theorem)$

اگر S یک آماره بسنده کامل برای θ و T یک آماره فرعی (آماره‌ای که توزیع آن به پارامتر وابسته نباشد) باشد، آنگاه S و T مستقل‌اند.

قضیه باسو $(Basu\ Theorem)$

اگر S یک آماره بسنده کامل برای θ و T یک آماره فرعی (آماره‌ای که توزیع آن به پارامتر وابسته نباشد) باشد، آنگاه S و T مستقل‌اند.

مثال ۱. اگر $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$ نشان دهید $\bar{X}$ و S^2 مستقل‌اند.

حل.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

می‌دانیم $\bar{X}$ آماره بسنده کامل برای μ می‌باشد. از طرفی S^2 آماره فرعی است، یعنی توزیع آماری آن به μ وابسته نیست. زیرا می‌دانیم:

$$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2),$$
$$\Rightarrow \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

لذا $\frac{(n-1)S^2}{1} \sim \chi^2_{(n-1)}$. لذا طبق قضیه باسو، $\bar{X}$ و S^2 مستقل‌اند.

مثال ۲. اگر $X_۱,\ldots,X_n \overset{iid}{\sim} N(\mu,۱)$, $\mu\in\mathbb{R}$. نشان دهید که $\bar{X}$ و $X_۱-\bar{X}$ مستقل‌اند.

حل. می‌دانیم که $\bar{X}$ آماره بسنده کامل برای μ است و نیز هر ترکیب خطی از متغیرهای نرمال دارای توزیع نرمال است. پس چون X_i ها مستقل و دارای توزیع نرمال هستند، توزیع $X_۱-\bar{X}$ نرمال است.

محاسبه میانگین و واریانس $X_۱-\bar{X}$:

$$E(X_۱-\bar{X}) = E(X_۱)-E(\bar{X}) = \mu-\mu = ۰,$$
$$Var(X_۱-\bar{X}) = Var(X_۱)+Var(\bar{X}) - ۲Cov(X_۱,\bar{X}).$$

داریم:

$$Cov(X_۱,\bar{X}) = Cov(X_۱, \frac{X_۱}{n}+\cdots+\frac{X_n}{n})$$
$$= Cov(X_۱,\frac{X_۱}{n}) + Cov(X_۱, \frac{X_۲}{n}+\cdots+\frac{X_n}{n}) \quad (\text{بسط کوواریانس})$$
$$= Cov(X_۱,\frac{X_۱}{n}) = \frac{۱}{n}Cov(X_۱,X_۱) = \frac{۱}{n}Var(X_۱) = \frac{۱}{n}. \quad (\text{مستقل بودن})$$

لذا $Var(X_۱-\bar{X}) = ۱+\frac{۱}{n}-\frac{۲}{n} = \frac{n-۱}{n}$.

پس $X_۱-\bar{X}\sim N(۰,\frac{n-۱}{n})$ و لذا $X_۱-\bar{X}$ آماره فرعی است. بنابراین از قضیه باسو آماره‌های $\bar{X}$ و $X_۱-\bar{X}$ از هم مستقل‌اند.

مثال ۳. اگر $X_1, \ldots, X_n \overset{iid}{\sim} Exp(\theta)$ با تابع چگالی زیر باشند

$$f(x,\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x > 0, \;\; \theta > 0,$$

و آماره‌های $S = \sum_{i=1}^{n} X_i$ و $T = \frac{X_1}{\sum_{i=1}^{n} X_i}$ تعریف شده باشند.

الف) نشان دهید S و T از هم مستقل‌اند.
ب) $E(T)$ را محاسبه کنید.

حل الف. می‌دانیم که آماره‌ی S برای θ بسنده کامل است (بنا بر خاصیت خانواده نمایی تک پارامتری).
حال باید نشان دهیم که T یک آماره فرعی است. برای این منظور قرار می‌دهیم $Y = \frac{X}{\theta}$. تابع چگالی Y با استفاده از تکنیک تبدیل متغیر به‌صورت زیر به‌دست می‌آید:

$$f_Y(y) = f_X(x)\,|\frac{dx}{dy}|.$$

$$y = \frac{x}{\theta}, \;\; \Rightarrow \;\; x = \theta\, y, \;\; \Rightarrow \;\; \frac{dx}{dy} = \theta,$$

$$f_Y(y) = f_X(\theta\, y)\,|\theta| = \frac{1}{\theta}\, e^{-\frac{\theta\, y}{\theta}}\, \theta = e^{-y}, \quad y > 0,$$

یعنی $Y \sim Exp(1)$ و توزیع Y به θ وابسته نیست.

داریم:

$$T = \frac{X_1}{\sum_{i=1}^{n} X_i} = \frac{X_1/\theta}{\sum_{i=1}^{n} X_i/\theta} = \frac{Y_1}{\sum_{i=1}^{n} Y_i}, \quad (\text{تابعی از } Y_1, \ldots, Y_n)$$

چون توزیع $Y_i,\ i = 1, \ldots, n$ وابسته به θ نیست، لذا توزیع T هم وابسته به θ نبوده و بنابراین T آماره فرعی است.

حل ب.

$$E(T) = E\left(\frac{X_1}{\sum_{i=1}^{n} X_i}\right) = ?$$

داریم:

$$T = \frac{X_1}{\sum_{i=1}^{n} X_i} = \frac{X_1}{S} \Rightarrow X_1 = T\,S$$

لذا

$$\begin{aligned} E(X_1) &= E(T\,S) \\ &= E(T)\,E(S) \quad (\text{چون T و S مستقل‌اند}) \end{aligned}$$

پس

$$E(T) = \frac{E(X_1)}{E(S)} = \frac{\theta}{n\,\theta} = \frac{1}{n}.$$

مثال ۴. فرض کنید $X_۱, \ldots, X_n \overset{iid}{\sim} U(\circ, \theta)$.

الف) نشان دهید $S = X_{(n)}$ و $T = \frac{X_{(۱)}}{X_{(n)}}$ مستقل‌اند.

ب) پیدا کنید $E[\frac{X_{(۱)}}{X_{(n)}}]$.

حل الف. می‌دانیم که $X_{(n)}$ یک آماره بسنده کامل است. نشان می‌دهیم که آماره $T = \frac{X_{(۱)}}{X_{(n)}}$ یک آماره فرعی است.

روش اول. با محاسبه تابع چگالی T می‌توان نشان داد به پارامتر θ بستگی ندارد.
روش دوم. بدون محاسبه تابع چگالی T نشان می‌دهیم که T یک آماره فرعی است:

$$T = \frac{X_{(۱)}}{X_{(n)}} = \frac{\min(X_۱, \ldots, X_n)}{\max(X_۱, \ldots, X_n)},$$

داریم:

$$X_i \sim U(\circ, \theta), \quad \Rightarrow \quad Y_i = \frac{X_i}{\theta} \sim U(\circ, ۱), \quad i = ۱, \ldots, n,$$

لذا داریم:

$$T = \frac{X_{(۱)}}{X_{(n)}} = \frac{X_{(۱)}/\theta}{X_{(n)}/\theta} = \frac{Y_{(۱)}}{Y_{(n)}}.$$

که تابعی از $Y_۱,\ldots,Y_n$ است. چون توزیع Y_i ها وابسته به θ نیست لذا توزیع T هم وابسته به θ نمی‌باشد. یعنی T یک آماره فرعی است. پس، از قضیه باسو، آماره‌های S و T مستقل‌اند.

حل ب.

$$T = \frac{X_{(۱)}}{X_{(n)}} = \frac{X_{(۱)}}{S} \;\Rightarrow\; X_{(۱)} = T.S.$$

از مستقل بودن T و S نتیجه می‌شود:

$$E(X_{(۱)}) = E(T)\,E(S),$$
$$\Rightarrow\; E(T) = \frac{E(X_{(۱)})}{E(S)} = \frac{E(X_{(۱)})}{E(X_{(n)})}.$$

داریم:

$$X_۱,\ldots,X_n \overset{iid}{\sim} U(۰,\,\theta), \quad \Rightarrow \frac{X_۱}{\theta},\ldots,\frac{X_n}{\theta} \overset{iid}{\sim} U(۰,\,۱),$$
$$\Rightarrow\; \frac{X_{(r)}}{\theta} \sim Beta(r,\,n-r+۱),$$
$$\Rightarrow\; E(\frac{X_{(r)}}{\theta}) = \frac{r}{r+n-r+۱} = \frac{r}{n+۱} \;\Rightarrow\; E(X_{(r)}) = \frac{r\,\theta}{n+۱}.$$

و لذا داریم:

$$E(T) = E(\frac{X_{(1)}}{X_{(n)}}) = \frac{\frac{\theta}{n+1}}{\frac{n\theta}{n+1}} = \frac{1}{n}.$$

تذکر. قضیه باسو را در مواردی که پارامتری در توزیع وجود نداشته باشد نیز می‌توان استفاده کرد.

مثال ۵. اگر $X_1, \ldots, X_n \overset{iid}{\sim} N(0, 1)$، نشان دهید که $\bar{X}$ و S^2 مستقل‌اند.

حل. موقتا فرض می‌کنیم که $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$. بنا بر مثال ۱، چون $\bar{X}$ یک آماره بسنده کامل برای μ است و S^2 یک آماره فرعی است، لذا $\bar{X}$ و S^2 مستقل می‌باشند برای هر $\mu \in \mathbb{R}$، از جمله برای $\mu = 0$.

مثال ۶. اگر $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ که σ^2 مجهول است. نشان دهید که $\bar{X}$ و S^2 مستقل‌اند.

حل. می‌دانیم برای هر σ^2 ثابت $\bar{X}$ یک آماره بسنده کامل برای μ و S^2 یک آماره فرعی است. لذا برای هر σ^2 ثابت $\bar{X}$ و S^2 مستقل‌اند و بنابراین با تغییر این σ^2 ثابت می‌توان نتیجه گرفت که $\bar{X}$ و S^2 برای هر σ^2 نیز مستقل‌اند.

تمرین ۱. اگر $X_1, \dots, X_n \overset{iid}{\sim} Exp(\frac{1}{\theta})$ با تابع چگالی زیر باشند

$$f(x,\theta) = \theta\, e^{-\theta\, x}, \quad x > 0, \ \ \theta > 0,$$

و آماره‌های $S = \sum_{i=1}^{n} X_i$ و $T = \frac{X_1}{\sum_{i=1}^{n} X_i}$ تعریف شده باشند.

الف) نشان دهید S و T از هم مستقل‌اند. ب) $E(S)$ و $E(T)$ را محاسبه کنید.

تمرین ۲. فرض کنید $X_1, \dots, X_n \overset{iid}{\sim} f(x,\theta)$ که

$$f(x,\theta) = \frac{2x}{\theta}, \quad 0 < x < \theta, \ \ \theta > 0.$$

الف) نشان دهید $T = X_{(n)}$ و $S = \frac{X_{(1)}}{X_{(n)}}$ از هم مستقل‌اند. ب) $E(S)$ و $E(T)$ را محاسبه کنید.

تمرین ۳. اگر $X_1, \dots, X_n \overset{iid}{\sim} f(x,\theta)$ که

$$f(x,\theta) = e^{-(x-\theta)}, \quad x > \theta, \ \ \theta > 0.$$

الف) نشان دهید $X_{(1)}$ و $\bar{X} - X_{(1)}$ مستقل‌اند. ب) $E[X_{(1)}(\bar{X} - X_{(1)})]$ را محاسبه کنید.